Let Rover Take Over : A Study of Mixed-Initiative Control for Remote Robotic Search and Detection

David J. Bruemmer, Douglas A. Few, Ronald L. Boring Julie L. Marble, Miles C. Walton
Human, Robotic and Remote Systems Department
Idaho National Laboratory
Idaho Falls, ID, USA
{bruedj, fewda, borir1, marbjl, mwalton}@inel.gov

Curtis Nielsen
Machine Intelligence, Learning, and Decisions Lab
Brigham Young University
Provo, Utah, USA
curtisn@cs.byu.edu

Abstract – *This article presents results from three successive experiments in which human operators were teamed with a mixed-initiative robot control system to accomplish various real-world search and detection tasks. By assessing human workload and error together with overall performance, these studies provide an objective means to contrast different modes of robot autonomy and to evaluate both the usability of the interface and the effectiveness of autonomous robot behavior. The first study compares the performance achieved when the robot takes initiative to support human driving with the opposite case when the human takes initiative to support autonomous robot driving. The utility of robot autonomy is shown through achievement of optimal performance when the robot is in the driver's seat; however, operators were sometimes confused by robot initiative and suffered from disorientation. In response, the second experiment introduces a virtual 3D map representation that supports shared understanding of the task and environment. When used to replace video, the 3D map reduces operator workload and error and, by drastically lowering bandwidth requirements, permits long-range, non-line-of-sight communication. The third experiment separates the various interface components into three workstations to assess the different roles demanded of the operator and how these roles change depending on the robot's level of autonomy. Results point to fundamental benefits of mixed-initiative control, showing that autonomous robot driving and decision making can increase performance and reduce error even when workload is distributed amongst multiple operators.*

Keywords: Human-robot interaction, usability evaluation

Introduction

Remote robotic operations often involve high operator workload, communication bandwidth constraints, long distances, and/or a lack of visible environmental features. Given these constraints, teleoperation seems a poor choice and yet almost all mobile ground systems in use today place responsibility for all decisions on the human. Despite the recognized need for autonomy, the performance of most intelligent robots does not approach that of a human operator and often fails to support operator trust. Unless the robustness of robot behavior and the flexibility of human-robot interaction methods improve, robots will continue to be excluded from many environments and tasks. Remote deployment of mobile robots offers one of the most compelling opportunities to merge human intelligence with machine proficiency. Rigorous, real-world HRI evaluations can illuminate the path towards this goal

Yanco, Drury, and Scholtz [1] have identified two major shortcomings in current HRI evaluations. The first is that the designers of the system are often used as test users. The second is that HRI evaluations are commonly informal and rarely provide controlled, objective assessment. Another shortcoming to previous human-robot interaction studies has been a lack in the number and diversity of participants. As suggested by Yanco et al., the present experiments do not use system designers or operators with extensive training. The focus of this research is not one particular application, but rather in the general question of how humans and robots can best cooperate; consequently, every effort was made to use a large and varied participant pool of novice users. The use of novice users maximizes both the relevance of the study to multiple applications and the evaluation's sensitivity to interface shortcomings. Of course, such an evaluation does not preclude the necessity of further evaluation with actual target users.

System Design and Implementation

Four robot modes of control are available from the interface [3, 4], affording the robot different types of behavior and levels of autonomy.

1. *Tele Mode* is a fully manual mode of operation, in which the operator must manually control all robot movement.
2. *Safe Mode* is similar to Tele Mode, in that robot movement is dependent on manual control. However, in Safe Mode, the robot is equipped with a level of initiative that prevents the operator from colliding with obstacles.
3. In *Shared Mode*, the robot relieves the operator from the burden of direct control using reactive navigation to find the optimal path based on the robot's perception of the local environment.
4. *Autonomous Mode* consists of series of high-level tasks such as patrol or search a selected region or follow a designated path. In Autonomous Mode, the only user intervention occurs on the tasking level; the robot itself manages all decision-making and navigation.

Robot Implementation

Operational experience in hazardous domains indicates that before it possible to explore the research questions pertaining to mixed-initiative control, it is essential to first develop trustworthy, effective robot behaviors. Since no single platform is appropriate for all tasks, the Idaho National Laboratory (INL) has developed a behavior architecture that can port seamlessly to a variety of robot geometries and sensor suites including those platforms in Figure 1. Experiments discussed in this paper utilized the "ATRV mini" (far left) and the "ATRV Jr" (center). On each robot, the behavior architecture utilizes a variety of sensor information including inertial sensors, compass, wheel encoders, laser, computer vision, thermal camera, infrared break beams, tilt sensors, bump sensors, sonar, and ultrasonic sensors.

Figure 1: The family of robots on which the INL control architecture resides

Interface Design

The default configuration of the interface consists of a single touch screen display containing five sizeable windows (see Figure 2). The upper left-hand window on the screen contains a video feed from the robot as well as controls for pan, tilt, zoom and changing camera settings such as toggling between thermal and visual imagery. Although the picture settings sub-window was not used by the participants, it allows operators to change the frame size, frame rate and compression on-the-fly to adapt to changing data link rates and support different task elements. The upper right-hand window contains status indicators and controls that allow the operator to monitor and reconfigure the robot's sensor suite as needed. The lower right-hand window pertains to the movement within the local environment and provides indications of direction and speed of robot motion, obstructions, resistance to motion, and feedback from contact sensors. At the far right, the user can select between different levels of robot autonomy. The lower central window provides an emerging map of the environment and allows the user to initiate a number of waypoint-based autonomous behaviors such as area search and follow path. The lower left-hand window contains information about the robot's operational status.

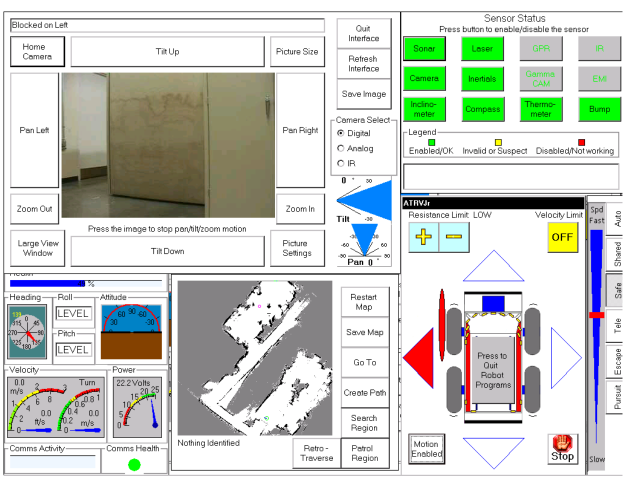

Figure 2: The standard interface configuration

When driving the robot directly, operators can give directional commands using the joystick. For each of the three experiments, participants were explicitly instructed on how to use the onscreen controls and the joystick. Depending on the interface configuration, operators may also pan, tilt and zoom the camera by using another, 3-axis joystick on the interface console.

Virtual 3D Display

The virtual 3D component has been developed by melding technologies from the INL [5], Brigham Young University (BYU) [6], and Stanford Research Institute (SRI) International [7,8]. The 3-D virtual display is not based on true 3-D range sensing, but rather by extruding a 2D map to provide the user with a malleable perspective. The INL control system uses a technique developed at SRI called Consistent Pose Estimation (CPE) that allows for efficient incorporation of new laser scan information into a growing map. Within this framework, SRI has found a solution to the challenging problem of loop closure: how to optimally register laser information when robot returns to an area previously explored (and ‘recognize’ that it was there previously). CPE provides an efficient means of generating a near-optimal solution to the constraint network, and yields high-quality maps.

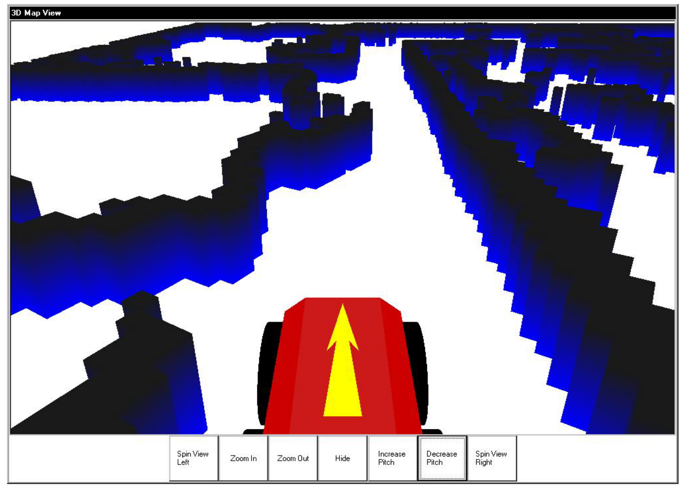

Figure 3: The virtual 3-D display

Note that the virtual 3D map may include not only obstacles, but also other semantic entities that are of significance to the operator. The operator may choose entities from a drop down menu or may insert translucent still images excerpted from the robot video. In this manner, the workspace can support both virtual and real elements, providing a means to remember what has been found and where. By changing the zoom, pitch, and yaw of the field of view, it is possible to move from an egocentric perspective (i.e. looking out from the robot), to a fully exocentric view where the entire environment can be seen at once. As Scholtz [9] points out, the roles of human operators do not remain static and interfaces should adapt accordingly.

Experiment 1

The goal of this study was to show that the interface was usable by a broad sampling of the population and that the robot behaviors developed for guarded motion and autonomous navigation were useful and effective. In particular, the goal was to see how autonomous driving would compare with direct joystick control.

Participants

The first study included 107 participants drawn as volunteers from attendees of the INL annual science and engineering exposition. The participants consisted of 46 females and 61 males, ranging in age from 3 to 78 years old, with a mean age of 14. It could be argued that attendees of a science and engineering exposition are likely to be more technologically savvy than the general populous. However when questioned, none of the participants had experience in remote system operations, thus qualifying them as novice users.

Procedure

A 20' x 30' robot search environment was created for this test. The participants controlled the robot from a remote station, thereby ensuring that they had no visual cues from the environment. To facilitate realistic maneuvering through an urban environment, the robot's search arena featured several obstacles. The central area was divided using conventional office dividers, while four cylindrical pylons were also placed strategically to force participants to maneuver effectively. Five objects were placed throughout the arena in fixed locations. These consisted of two mannequins, a stuffed dog, a disabled robot, and a small, simulated explosive device. The placement of these items also made the actual driving task more challenging.

Participants were given 60 seconds to locate as many of the five items in the search area as possible. Each participant was instructed on the use of the joystick for controlling the robot. Additionally, each participant was instructed on the robot's camera controls (e.g. pan, tilt, zoom). Runs alternated between use of Safe Mode where the robot takes initiative only to protect itself from collisions and Shared Mode where the robot drives autonomously, but accepts periodic intervention from the operator. For participants using Shared Mode, it was explained that they should let the robot do the driving, but that if they wanted to redirect the robot, the robot would temporarily yield control to their joystick commands. There was no need to experimentally compare either Shared Mode or Safe Mode to Tele Mode. Although Tele Mode can be useful for expert users to push open a door or shift a chair out of the way, observation was sufficient to show that allowing novices to remotely operate 200lb robots without guarded motion was both remarkably inefficient and acutely dangerous.

Results

The effects of participant age, gender, and operational mode were compared against the total number of objects that were located and identified (see Figure 4). In contrast to a previous experiment [2] which emphasized subjective measures such as trust, ease-of-use and feeling of control, this experiment focused on a quantitative assessment, using the number of items found as the performance metric. The results were analyzed by age in five-year intervals up to 20 years old; thereafter they were grouped in ten-year intervals. This ensured that the analysis was sensitive to possible developmental differences in pre-adults. There was no significant difference in the number of objects found across participants of different ages, F (8, 96) = 1.64, p=0.12. Although the age group 15-20 had the highest overall average, analysis of the data showed that the fluctuations in the number of objects found by different age groups were not meaningful. There was no difference in the number of objects found due to gender. Females statistically found the same number of objects as did males, M=2.54 and M=2.68 respectively, F (1, 103) = 0.31, p=0.58. There was a significant difference due to operational mode, F (1, 103) = 4.83, p<0.05. Participants who used Shared Mode found an average of 2.87 objects, while those who used Safe Mode found an average of 2.35 objects. There were no significant twoway or threeway interactions between gender, age, and operational mode.

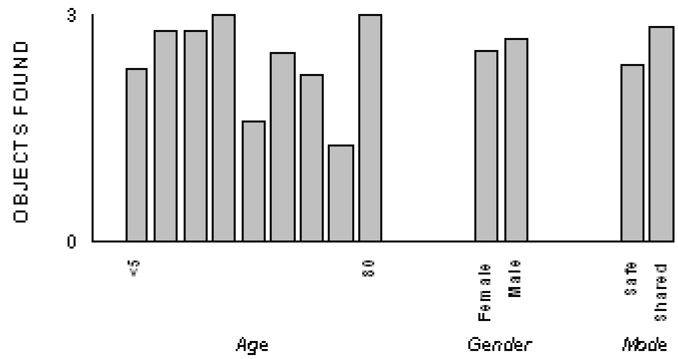

Figure 4: The average number of objects (out of five)

Discussion

The interface proved to be highly usable by novices, regardless of age or gender. Encouragingly, participants met with success in both Safe Mode and Shared Mode, indicating that the guarded motion and autonomous navigation behaviors were effective and useful. Performance was significantly better in Shared Mode than in Safe Mode, suggesting that navigational autonomy can enhance remote operations by freeing the user to focus on the application-oriented aspects of the task. Subjective assessment indicated that most participants felt in control, but in quite a few cases, participants were confused by robot initiative and engaged in a fight for control. It was observed that in most of these cases, the operators seemed to be focused entirely on the video and failed to notice the textual and graphical indications from the robot that movement was obstructed in the direction demanded by the operator. Also, some participants became disoriented and failed to remember where they had already been.

Although the study demonstrates the utility of robot autonomy, it leaves many questions unanswered. Beyond looking at overall performance (e.g. items found) the first experiment fails to discern the reasons for the observed difference in performance. In order to nuance the differences between guarded-teleoperation and autonomous robot navigation, it is necessary to empirically measure differences in operator workload, operator error, and operator confusion. Also, this experiment utilized a relatively small environment. Areas of the environment required careful maneuvering, but the task was not designed to reward path planning or strategy. Likewise, although some operators experienced temporary disorientation, there was little possibility of the operator getting truly lost. The question was raised of whether, in a more complex environment where intelligence is necessary, the robot's ability to make decisions and navigate autonomously would fall short of the humans. Finally, there was the question of whether robot autonomy was useful only because the operators were also faced with the task of looking at the video to locate and identify objects.

Experiment 2

The second and third experiments were designed with these questions in mind. The first experiment had shown that video was not sufficient to provide users with an understanding even of the local environment, much less the global environment. There is no doubt that humans are visually centric and generally prefer pictures and diagrams when attempting to understand or communicate. However, the first experiment raised the question of how useful streaming video actually was to users. Especially in tight spaces, where situation awareness is most important, the entire visual field may be filled by an immediate obstacle; conversely, the visual field may fail to show an obstacle if it is outside of the current visual field. In such cases, video communicates very little about the environment and can promote a false sense of situation awareness, which can lead to operator confusion and a fight for control. In response to this observation, work began to develop a new interface component that could appeal to the visually-centric human operator, while providing an effective means to represent the environment and communicate about the task.

Video demands high-bandwidth, continuous communication and is therefore ill-suited for many of the very environments where robots are most needed. Except for short ranges (< 100 meters), transmission of high-bandwidth video is only possible when line of sight can be maintained either with a satellite or another radio antenna. For instance, high-bandwidth video cannot be transmitted through layers of concrete and rebar, making it inappropriate for urban terrain. Likewise, forest and jungle canopy precludes reliable transmission of video. It has long been assumed that advances in communication will one day alleviate these technical limitations, but at the present time, reliable transmission of streaming video remains an elusive goal.

Drawing from the world of computer gaming, a virtual 3D display (see Figure 3) was developed as a means to give users insight into the reason for robot initiative as well as to diminish the possibility of disorientation. The second experiment was designed to empirically evaluate the effectiveness of the virtual 3D display, especially in contrast to reliable streaming video display. Unlike the first experiment that examined the difference between Shared and Safe Modes, the second experiment used only Safe Mode in order to insure that control mode did not complicate the analysis of the virtual 3D display.

Participants

The experiment was performed over a seven-day period within the St. Louis Science Center and utilized 64 visitors who volunteered to take part in the experiment. Participants ranged in age from 12 to 60 with most being junior high or high school students from schools in the St. Louis area.

Procedure

As before, the experiment was set up as a true remote deployment such that the operator control station was located several stories above and several hundred feet to the side of the arena where the robot operated. This arena was built by the production staff of the Science Center and utilized fake rocks, fake trees and mannequins as well as plywood dividers to create a maze environment (see Figure 5). In order to make the comparison between video and the 3D map representation conclusive, every effort was made to provide the best video possible. Trial runs indicated

Figure 5: A partial view of the arena built at the St. Louis Science Center

that video participants would be at a significant disadvantage simply because the ambient lighting, although normal, cast shadows that made it difficult to navigate with video. Although this is often the case in real-world deployments, the production staff augmented the ceiling with additional lighting that provided uniform lighting throughout the environment.

Due to the distance and physical occlusions separating the control station from the actual robot environment, analog video was not an option. Instead state of the art video compression was used to digitize the analog video into an MJPEG format and efficiently transmit from the robot to a wireless access point connected to the building's network. Exploiting the wired infrastructure in place throughout the building made it possible to provide continuous, reliable video, which far exceeded the performance possible through a wireless data link.

Figure 6: A near-complete map built by one of the participants.

Each participant was given basic instructions on how to use the interface, and, as with the other two experiments, no participant was permitted to drive the robot until the start of the trial run. Each trial run was exactly 3 minutes. At the beginning of each run, the robot was reset in order to erase the map. Each participant was told to direct the robot around the environment in order to build as big a map as possible. This task involved spatial reasoning because the operator must perceive the frontiers of the map and direct the robot to them in an optimal fashion. All participants were given access to the same 2-D map component (see Figure 6) within which a map emerges as the robot explores new ground. Exactly one half of the participants used the interface as depicted in Figure 2. These participants were able to use both the 2-D map and the video module. For the other half, the virtual 3-D interface module entirely occluded the video module.

During each trial, the interface stored a variety of useful information about the participant's interactions with the interface. Joystick bandwidth was recorded as the number of messages sent from the joystick indicating a change of more than 10% in the position of the stick. The interface also records the number of times that the robot was forced to take initiative to prevent a collision. It is important to note that the robot will only take initiative at the last possible moment that it can safely avoid a collision. Regardless of the robot's rotational or translational velocity, the robot always comes to a stop approximately two inches from the obstacle. The interface indicates physical blockages that impede motion in a given direction as red ovals next to the iconographic representation of the robot (lower left of figure 2). When the robot takes initiative to stop, the user should be able to discern that the robot is blocked based on these indications. However, previous experiments taught us that not every operator is able to attend to these visual indications. As a result, a force feedback joystick was implemented to resist movement in the blocked direction. Once the robot has already taken initiative to stop, if the human fails to understand the situation and tries to advance the robot into an obstacle, the joystick vibrates and emits a loud noise. It is specifically these instances which the system automatically logs as human error. Consequently, the metric referred to as human error, indicates not only human error, but also human confusion. For each trial the map produced by the robot was saved. In order to metric performance, a software algorithm was implemented to calculate the percentage of the full map that was present in each of these saved maps. This approach provides a reasonable, objective assessment and is much more relevant than a measure of distance traveled.

Results

This experiment focused on a quantitative analysis of performance, workload, error, and feeling of control collected during the exploration task. Based on the results of the first experiment it was assumed no statistical performance differences exist across age or gender differences. In the three minutes provided, the majority of participants explored over 50% of the total environment. Only one person, a 3D display participant, was able to build the entire map in the allotted 3 minutes.

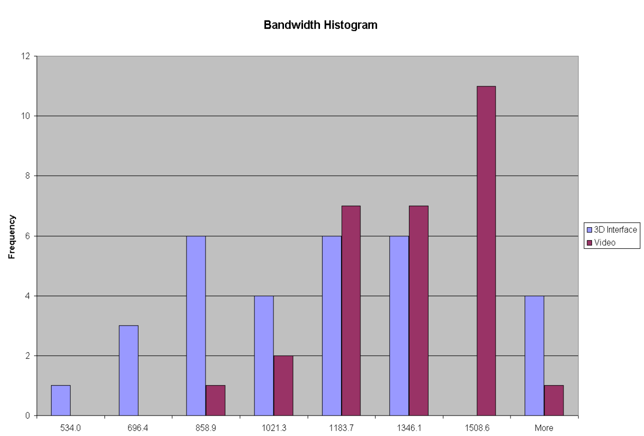

Figure 7: Joystick bandwidth histogram

As described above, task performance was calculated by comparing the map generated during the exploration task with the complete map of the task environment. This comparison showed no significant statistical difference between the use of the video interface module and the virtual 3D map module, *M* .71, *M* .61, respectively, $F(1, 31) = 0.558$ $p = 0.070$.

Using joystick bandwidth as an indication of human workload and robot initiative as a metric for human error, analysis shows that operators using the virtual 3-D display worked less and demonstrated fewer instances of navigational error. On average the joystick bandwidth for participants using the virtual 3D display was 1057.50 messages from the interface to the robot, compared to 1229.07 for operators using video feed, $F(1, 31) = 2.024$, $p < 0.05$. The robot initiative for participants using the virtual 3D display averaged 11.00, compared to an average of 14.29 for the video participants, $F(1, 31) = 0.399$, $p < 0.05$.

In addition to reduced workload and fewer errors, use of the virtual 3D display slightly increased the operator's subjective "feeling of control" while operating the robot. The average feeling of control for the 3D display was 7.219 compared with an average of 7.059 for the video, $F(1.31) = 0.497$, $p < 0.05$.

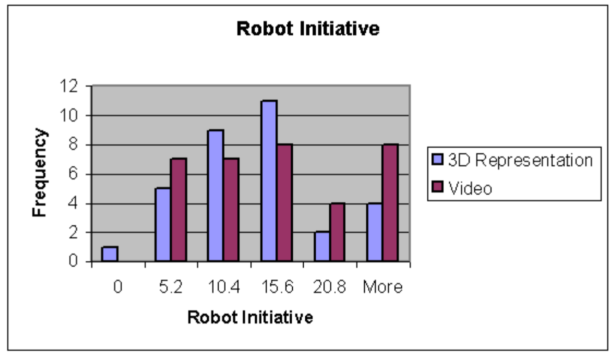

Figure 8: Robot Initiative Histogram

Discussion

This experiment provides compelling evidence that it is possible to support the visually centric needs of human operators without using video. Further experiments will be required to determine conclusively if there is any performance difference between the abstracted representation and the live video. It is important to note that since the St. Louis experiments, several significant improvements have been made to the virtual 3D display including refresh rate which cannot help but improve the utility of the virtual 3D interface. In its original instantiation, the virtual 3D display incurred no significant loss of performance and provided a reduced workload, fewer errors, and a heightened sense of control.

The abstracted data necessary to produce the virtual 3D representation can be sent over a low-bandwidth data transmission such as a single cell-phone or a long range radio. Whereas the video alone required at least 3,000,000 bits / second, the total interface bandwidth while using the virtual 3D interface was only 64,000 bits / second. This bandwidth savings allows control to extend into new domains. Work has already begun to apply the same control system used in this experiment to exploration of underground bunkers, caves, nuclear reactors, and urban search and rescue sites.

However, the fact that the human-robot team can function without video, is no reason to disregard the potential benefits of video in those instances when video is available. Experience with operators and subject area experts from Energy, Defense and Emergency Management contexts indicate that operators expect and can exploit video in remarkable ways. After all, many applications require the human to play a role in visual search and detection. Although the second experiment showed that video could be replaced with the 3D representation, the optimal interface will likely provide a dynamic balance between the video and virtual displays.

Experiment 3

The second experiment provided initial validation for the 3D map representation and indicated progress towards addressing the issues of disorientation and operator confusion observed in the first experiment. The third experiment was designed to revisit the comparison between modes of control. Would the virtual 3D display make the difference between modes more or less pronounced? How would the utility of this new component change depending on the control mode of operator control?

The third experiment was also intended to explore how the interface should be configured to support specific operator functions: navigation which depends on an exocentric display, driving which uses an egocentric display, and operation of an application payload which can be controlled independently from the robot. By assigning members of a team these different roles, the intent was to provide insight regarding the interdependence of these operator roles. The hypothesis was that the navigators would emerge as the leaders, but that conflict resolution and initiative from the drivers would also play a major role.

Another objective of the third experiment was to show that the benefits of using robot autonomy demonstrated in the first experiment were not merely due to the high cognitive workload placed on the operator. The typical assumption found in the literature is that robot autonomy trails behind human performance, but becomes increasingly useful as operator workload increases or communications fail [11,12,13]. While it is true that autonomy can support changing operator workload and communication constraints, it was hoped that this experiment could move beyond this assumption to show that even if optimal communication is maintained and human workload is kept at a minimum, the robot is still the better driver.

Participants

This experiment included 120 volunteers grouped into teams of six. Unlike the volunteers who comprised the participant pool for the last experiment, these participants signed up in advance to take part in the study and knew the other people in their group already. The vast majority of participants were high-school age, although several adults also participated.

Procedure

The experiment was run over seven days at the St. Louis Science Center. As with the previous experiment, the robot was located in the lower level of the Science Center while the control center was located on the top level. The same environment was used with the same lighting and placement of obstacles. Three mannequins were placed in hard-to-reach locations and remained in place throughout the entire experiment. The starting point of the robot alternated between two different locations such that an equal number of shared mode and safe mode runs were begun from each starting point.

For this experiment, the control interface components were divided amongst three separate stations, each with its own monitor and input devices. No interface component was visible at more than one control station. Two participants manned each station resulting in a total of six people dedicated to robotic system control. Having two participants at each station was not necessary, but insured that workload was minimal. For instance, one driver used the joystick while the other monitored the local environment window and activated the motion brake to halt autonomous driving when necessary. The stations were arranged in an arc such that the participants at each station could communicate easily with the others, but could not see the other displays.

The first control station was dedicated to the application payload, which in this case was a pan-tilt-zoom camera. Using a joystick that allowed them to operate the various camera controls, the application payload participants used the visual feedback from the robot to seek out the three mannequins and to provide navigational advice. The second control station was dedicated to driving the robot. Participants were permitted to see the virtual 3D window, the local environment window, the sensor status window and the robot state window (see Figure 2). Primarily, the operators at the driving station used the 3D virtual display, but were constrained to an egocentric perspective which precluded a global view of the environment. The final station was the navigation station. The navigators had access to the 2D map being built as the robot traveled through its environment which gave them a god's eye view of the environment and the robot's position in it. In addition, the participants at the navigation station were given an *a priori* map that showed the locations of the 3 mannequins. Task completion required the three groups to self-organize in order to arrive at and gain a visual lock on all three of the mannequins as quickly as possible. As in the previous experiment, joystick usage was measured as an indication of operator workload. Also, the force feedback response from the joystick was recorded as an indication of operator error and confusion.

Results

On average, less time was required for the participants using the higher level of robot autonomy. The mean completion time for Shared Mode participants was 466.8 seconds compared to the 641.1 second average completion time for the Safe Mode participants, $F\,(1, 9) = 3.64$, $p < 0.05$.

Safe Mode participants experienced greater workload than that of their Share Mode counterparts, $M = 2743.8$ and $M = 1725.6$ respectively, $F\,(1, 9) = 0.296$, $p < 0.05$. Using joystick vibration as a metric for human error shows that Safe Mode participants made 25.1 errors compared to 16.8 errors for the Shared Mode participants. $F\,(1, 9) = 5.92$, $p < 0.05$.

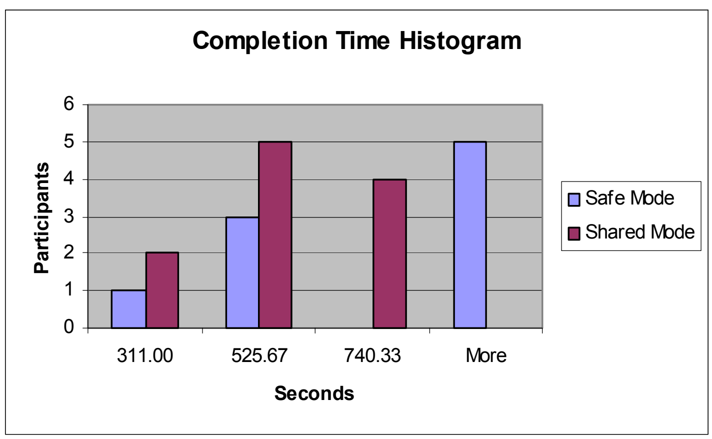

Figure 9 : Time to complete the task in shared and safe modes.

In terms of overall error and workload, the differences between Shared and Safe Mode are not independent of the difference in time. It stands to reason that the more time spent on a task, the more work will have to be done and the greater the probability of errors. An analysis of Safe Mode indicates that both error and workload were significantly correlated to task duration $r(7) = 0.761$, $p>0.05$ and $r(7) = 0.729$, $p>0.05$ respectively. Interestingly, Shared Mode only indicates a loose correlation between error and task duration $r(9) = 0.659$, $p>0.05$ and no correlation between task duration and human workload $r(9) = 0.311$, $p<0.05$.

Discussion

As with the first experiment, Shared Mode participants enjoyed increased performance efficiency and decreased workload when compared to their Safe Mode counterparts. Interestingly, the third experiment presents us with a more pronounced disparity between the two modes than the first. This result brings us to reject the hypothesis that decreasing individual human workload and increasing the role of strategy and intelligence would diminish the advantages of autonomous robot decision making and driving. The fact that the only substantial change to the interface between experiment one and three was the addition of the virtual 3D map, suggests that the ability of this component to support a shared cognitive understanding between the human and robotic team member is responsible for augmenting the advantages of shared control.

An interesting observation was that some of the operators responsible for driving the robot came to trust robot suggestions and robot initiative over that of their human team members. For many of the teams that used Shared Control, the robot emerged as a valued and trusted team-member. On the other hand, the operators often chose to override robot initiative, just as they sometimes chose to ignore advice from the navigators or payload operators. The key observation is that in all cases, team success depends on understanding the capabilities and limitations of each team member's role. For instance, in some teams, the navigator using the 2D map would give commands such as "Go to the top of the map" which meant nothing to the operators responsible for driving since they were using an ego-centric perspective. Once each workstation team understood the cognitive perspective of the other team members, a great deal more could be accomplished.

In many operational scenarios, it is not only possible, but probable that the roles of driving, navigating and operating the application payload will be spread amongst multiple human operators. This experiment yielded several observations that can inform such a scenario. For instance, successful teams included a navigator who took charge, especially once a mapping could be made from the a priori map to the emerging 2D map. Just as performance can be degraded by a fight for control between the driver and robot, there were similar instances of a fight for control between human functions. Effective teams communicated their level of confidence. In the beginning of a run, the navigators had less map built and therefore their advice was less certain and less useful. Communicating uncertainty helped the drivers arbitrate between advice from the navigator and payload stations. In a fielded system, a division of labor amongst multiple roles would not require team members to view only their own display, which would alleviate confusion and achieve resolution in the event of a conflict.

Conclusions

For several decades it has been assumed that humans should interact with robots primarily through a master-slave relationship based on streaming video sent from the robot to the human operator. The experiments discussed indicate that mixed-initiative control can provide a compelling alternative. Across a variety of homeland defense, military, department of energy, space exploration and industrial contexts, it is possible to apply this new interaction method to a broad range of tasks and applications, especially those where continuous video is not possible. Interfaces built around video are appropriate primarily for a master-slave relationship and are unsuitable for monitoring dynamic autonomy systems that permit different levels of operator involvement.. In contrast, abstracted representations such as the virtual 3-D display can promote dynamic autonomy and allow the potential benefits of mixed-initiative control to be more fully realized. Unlike video, which offers only a first person perspective, the 3D display adapts to changing levels of operator involvement and autonomy.

This advantage is especially important for multiple robot operations where it becomes impossible for a single operator to monitor or task multiple robots in a teleoperated fashion. Current work at the INL is adapting the virtual 3D display for use in countermine operations where multiple robotic vehicles used for demining can contribute to and be tasked via the same display. A future study will use this multi-robot scenario to investigate the value of Shared Mode when compared against Autonomous Mode. Research thus far has shown that when accompanied by an interface that supports shared understanding and control, robot initiative can be used to reduce human workload, increase overall performance and enable new robotic applications.

References

[1] H. A. Yanco, J. L. Drury, and J. Scholtz, "Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition," *Journal of Human-Computer Interaction*, Vol. 19, pp. 117-149, 2004.

[2] J. L. Marble, D. J. Bruemmer, and D. A. Few, "Lessons learned from usability tests with a collaborative cognitive workspace for human-robot teams," *IEEE Conf. on Systems, Man, and Cybernetics*, 2003.

[3] D. J. Bruemmer, J. L. Marble, and D. D. Dudenhoeffer, "Mutual Initiative in Human-Machine Teams," *Proc. IEEE Conference on Human Factors and Power Plants*, pp. 7.22-30, 2002.

[4] D. J. Bruemmer, J. L. Marble, D. D. Dudenhoeffer, Intelligent Robots for Use in Hazardous Environments, *PerMIS 2002*, Gaithersburg, MD, August 2002.

[5] D. J. Bruemmer, D. A. Few, R. Boring, M. Walton, J. L. Marble, C. Nielsen, J. Garner. "'Turn Off the Television !:' Robotic Exploration Experiments with a 3-D Abstracted Map Interface ," In *Proceedings of the 38th Hawaii International Cconference on the System Sciences,* Waikoloa Village, Hawaii, January 2005 .

[6] C. Nielsen, B. Ricks, M. Goodrich. D. Bruemmer, D. Few, M. Walton. "Snapshots for Semantic Maps," In *Proceedings of Systems, Man and Cybernetics 2004*, The Hague, Netherlands, October 10-13, 2004.

[7] K. Konolige. Large-scale map-making. *In Proc. of AAAI,* San Jose, CA, 2004.

[8] J.S Gutman and K. Konolige, "Incremental Mapping of Large Cyclic Environments", *CIRCA 99,* Monterey, California, 1999.

[9] J. Scholtz, Human-Robot Interactions: Creating Synergistic Cyber Forces, *In Proceedings from the 2002 Workshop on Multi-Robot Systems*, Washington, D. C. March, 2002.

[10] D. Bruemmer, R. Boring, D. Few, J. Marble, M. Walton. "I Call Shotgun!: An Evaluation of Mixed-Initiative Control for Novice Users of a Search and Rescue Robot," In *Proceedings of Systems, Man and Cybernetics 2004*, The Hague, Netherlands, October 10-13, 2004.

[11] C. Nielsen, M. Goodrich, J. Crandall. "Experiments in Human-Robot Teams," In *Proceedings from the 2003 International Workshop on Multi-Robot Systems*, Washington D.C., March 2004.

[12] B. Trouvain, H. L.Wolf, F. E. Schneider, "Impact of Autonomy in Multi-Robot Systems on Teleoperation Performance," In *Proc. of the 2003 Workshop on Multi-Robot Systems*, Washington D.C., March 2004.

[13] M. A. Goodrich, D. R. Olsen Jr, J. W. Crandall, and T.J. Plamer. "Experiments in Adjustable Autonomy." In *Proc. of IJCAI Workshop on Autonomy, Delegation, and Control : Interaction with Autonomous Agents,* Seattle, Washington, August 2001.